



EX PARTE OR LATE FILED

Jay C. Keithley
Vice President

Law & External Affairs
1850 M Street, NW, Suite 1100
Washington, DC 20036
Voice 202 828 7453
Fax 202 296 3469
jav.c.keithley@mail.sprint.com

RECEIVED

APR 23 1998

FEDERAL COMMUNICATIONS COMMISSION
OFFICE OF THE SECRETARY

EX PARTE

April 23, 1998

Ms. Magalie Romas Salas
Secretary
Federal Communications Commission
1919 M Street, N.W. Room 222
Washington, D.C. 20554

DOCKET FILE COPY ORIGINAL

RE: In the Matter of Federal-State Joint Board on Universal Service, CC Docket No. 96-45
and Forward-Looking Mechanism for Non-Rural LECs, CC Docket No. 97-160

Dear Ms. Salas:

The attached letter was today provided to Mr. A. Richard Metzger, Jr., Chief,
Common Carrier Bureau in regards to the above referenced dockets.

Sprint requests that this information be made a part of the record in this matter. Two
copies of this letter, in accordance with Section 1.1206(a)(1), are provided for this purpose. If
you have any questions, please feel free to call.

Sincerely,

Jay C. Keithley

Attachment

of Copies rec'd 0+3
CODE



Jay C. Keithley
Vice President

Law & External Affairs
1850 M Street, NW, Suite 1100
Washington, DC 20036
Voice 202 828 7453
Fax 202 296 3469
jay.c.keithley@mail.sprint.com

April 23, 1998

Mr. A. Richard Metzger, Jr.
Chief, Common Carrier Bureau
Federal Communications Commission
1919 M Street, NW, Room 500
Washington, DC 20554

RE: In the Matter of Federal-State Joint Board on Universal Service, CC Docket 96-45
and Forward-Looking Mechanism for Non-Rural LECs, CC Docket 97-160

Dear Mr. Metzger:

Earlier this week and last Thursday, Sprint met with several members of your staff, as well as the Common Carrier Legal Assistants copied below, to describe a potentially fatal flaw in the HAI cost proxy model. As described in the ex partes filed in connection with the above-mentioned meetings and below, late last week, during a Nevada Public Utilities Commission-ordered examination of some of the underlying data used in the HAI cost proxy model, Sprint discovered that, for Nevada (the only state for which the model sponsors have provided access to the data), the HAI model underestimates distribution plant investment by a factor of nine. Based on the Nevada data, the flaw is neither an input issue nor a matter of customer location; the flaw is systematic. The Nevada data (not to mention the sponsors refusal to provide greater access to the relevant data) strongly indicates that the HAI model fails to meet the standards the Commission has ruled must be met before a cost proxy model can be used in the development of the Commission's new high cost universal service support mechanism. This development demands that the HAI model be subjected to additional analysis – by the Commission and by others, including Sprint. Accordingly, as set out more specifically below, Sprint respectfully asks the Bureau to order the HAI model sponsors to provide access to “all underlying data, formulae, computations, and software associated with the model,” including all its pre-processing algorithms, with associated raw data.

In its May 8, 1997 Order in Docket 96-45, the Commission set forth a list of ten criteria with which any forward looking cost model submitted for Commission consideration must comply. The eighth criteria states:

“[t]he cost study or model and all underlying data, formulae, computations, and software associated with the model should be available

to all interested parties for review and comment. All underlying data should be verifiable, engineering assumptions reasonable, and outputs plausible."

The HAI cost model makes use of proprietary data (geocoded customer location data) in the formulation of its clusters which are the major component of the distribution plant module of that model. In the HAI model's preprocessing, customers are located using geocoded data; non-geocoded customers (referred to as surrogate points) are placed on the perimeters of census blocks; the model then uses a mathematical algorithm to transform that customer location data into "clusters," which become the basis for calculating the amount of distribution plant required to serve those customers. Only the summary results of these calculations are made available in the model.

As a result, no interested party has had the capability of reviewing the accuracy or plausibility of the manifold assumptions and calculations made by the HAI model in determining the required amount of distribution plant. The HAI sponsors have contended that this data is proprietary and have not permitted Sprint, or any other interested party, including the Commission staff, access to the data and calculations underlying its distribution plant module.

Recently, the Nevada Public Service Commission required AT&T to permit parties involved in a cost proceeding in that state access to this underlying data, subject to a nondisclosure agreement that would protect the confidentiality of the proprietary geocoded customer location data. On April 16, Sprint was able, *for the first time*, to review and analyze the calculations underlying the HAI model's determination of distribution plant. Sprint was able to review data only for the state of Nevada.

Sprint's analysis has revealed a systematic and significant bias in the HAI distribution plant module, the effect of which is to understate distribution plant distances and, therefore, distribution plant costs. Based on a sample of clusters analyzed by Sprint, the HAI model understated the amount of distribution plant required by a factor of nine. This factor was derived by simply comparing the minimum distance required to connect the actual geocoded and surrogate customer locations (data which was produced by HAI, not Sprint) with the distribution plant distances calculated by the HAI model. It is important to note that understatement occurs whether the cluster points are actually geocoded or whether surrogate points are used. It is clear, then, that this is not a geocoding problem; rather, it is a fundamental problem in the HAI model platform. A detailed explanation of the flaws in the HAI algorithms which produce this systematic understatement of distribution plant requirements was provided to Commission staff, and is also attached to this letter.

Unequivocally, it is not plausible that the amount of distribution plant needed to serve a group of customers is less than 1/9 of the actual physical distance between those

customers. This systematic understatement of distribution plant requirements, particularly for rural, less densely populated areas, has serious ramifications for the determination of Universal Service Fund support. Assuming that the HAI model understates distribution plant by a factor of from 2 to 5—and not by the factor actually found for the clusters Sprint was able to examine-- correcting that understatement could produce increases in USF funding requirements of 60% to 250%.

Obviously, this is only an estimate, since Sprint does not have access to the data necessary to reliably quantify the impacts of this flaw in the HAI model on a nationwide basis. At this point, the only firm conclusion that can be reached is that the HAI model systematically understates distribution plant to such a degree that the outputs of that model are not plausible. A determination of the precise magnitude of that understatement cannot be made unless the HAI sponsors are required to provide access to all data, for all states.

Sprint therefore requests your assistance in obtaining access to this data. Specifically, Sprint requests that it, and all other interested parties, be permitted access to the data and all associated calculations and algorithms in order to perform the following calculations:

1) the diagonal length of the minimum bounding rectangle of the original HAI cluster. (Logically, the diagonal is the theoretically absolute minimum length of distribution cable required to serve the identified customers);¹

2) the minimum spanning tree for each cluster. (The minimum spanning tree is a standard mathematical measure of the minimum distance needed to connect any set of points);² and

3) the percentage of customer locations actually geocoded for each cluster.

Sprint is not requesting that any of the underlying proprietary data in the HAI model be made public. The information requested herein divulges neither the identity nor the location of the customers-the data which the HAI sponsors claim to be proprietary to

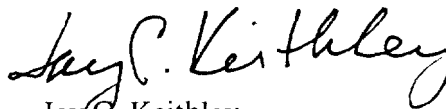
¹ In the vast majority of cases, the diagonal length will still significantly understate the amount of distribution cable required. The only case in which the diagonal would represent the actual required cable is when all of the customers in a cluster are actually located in a straight line coinciding with that diagonal.

² Sprint recognizes that, in some instances, the minimum spanning tree will overstate distribution lengths relative to a "star" configuration. Although it believes any such overstatement would be minimal (in the area of 10% at most), Sprint is working to develop mathematical algorithms to compute distribution distances utilizing a star configuration, and to the extent it was able to do so, would also perform that analysis on the HAI geocoded and surrogate points.

Mr. A. Richard Metzger, Jr.
April 23, 1998
Page 4 of 4

those who claim proprietary interest in the data. Further, the analyses Sprint requests it be permitted to perform on HAI distribution module is both consistent with the Commission's own criteria, cited above, and absolutely essential to a determination of the reasonableness and plausibility of the results produced by the HAI model.

Respectfully,

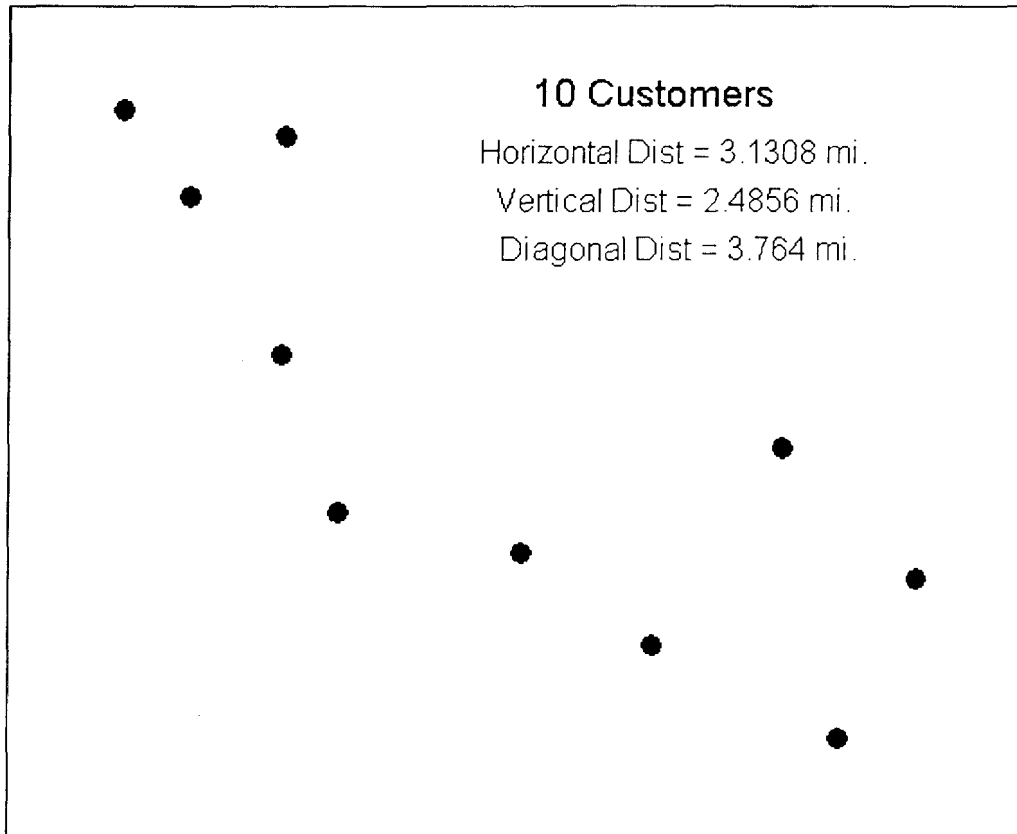

Jay C. Keithley

Attachment

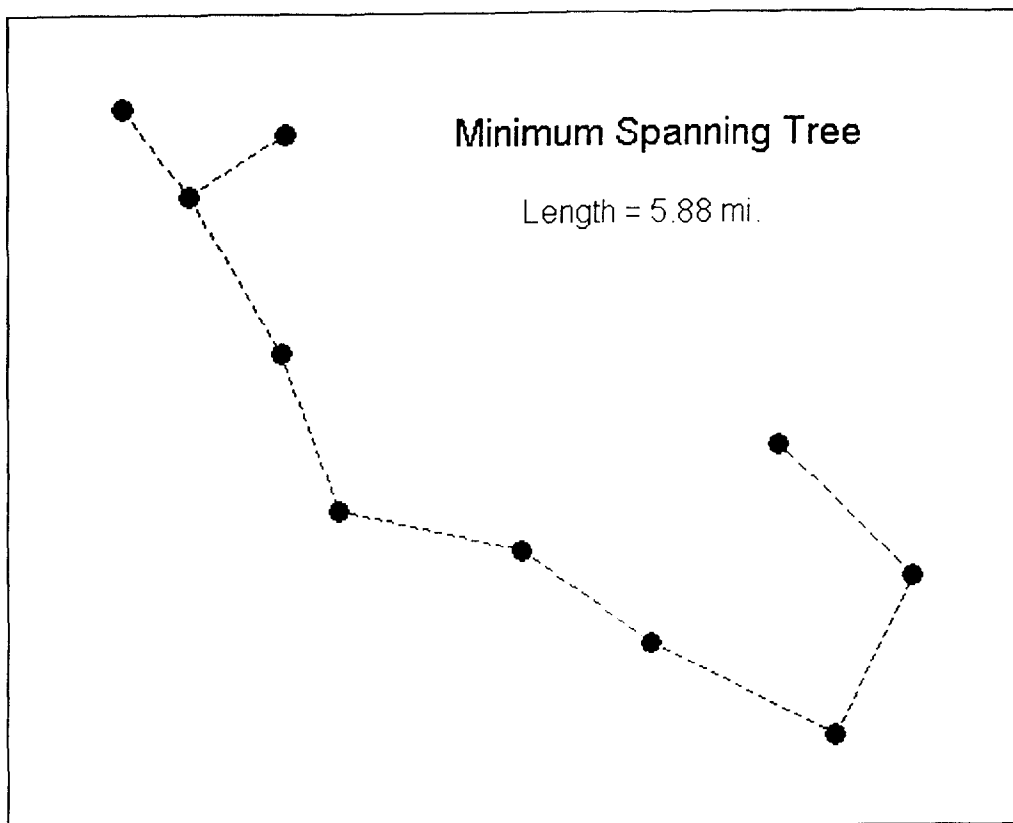
cc: Chairman William Kennard
Commissioner Susan Ness
Commissioner Gloria Tristani
Commissioner Michael Powell
Commissioner Harold Furchtgott-Roth
Thomas Power
Jim Casserly
Paul Gallant
Kyle Dixon
James B. Schlichting

Hatfield's Polygons Converted to Rectangles

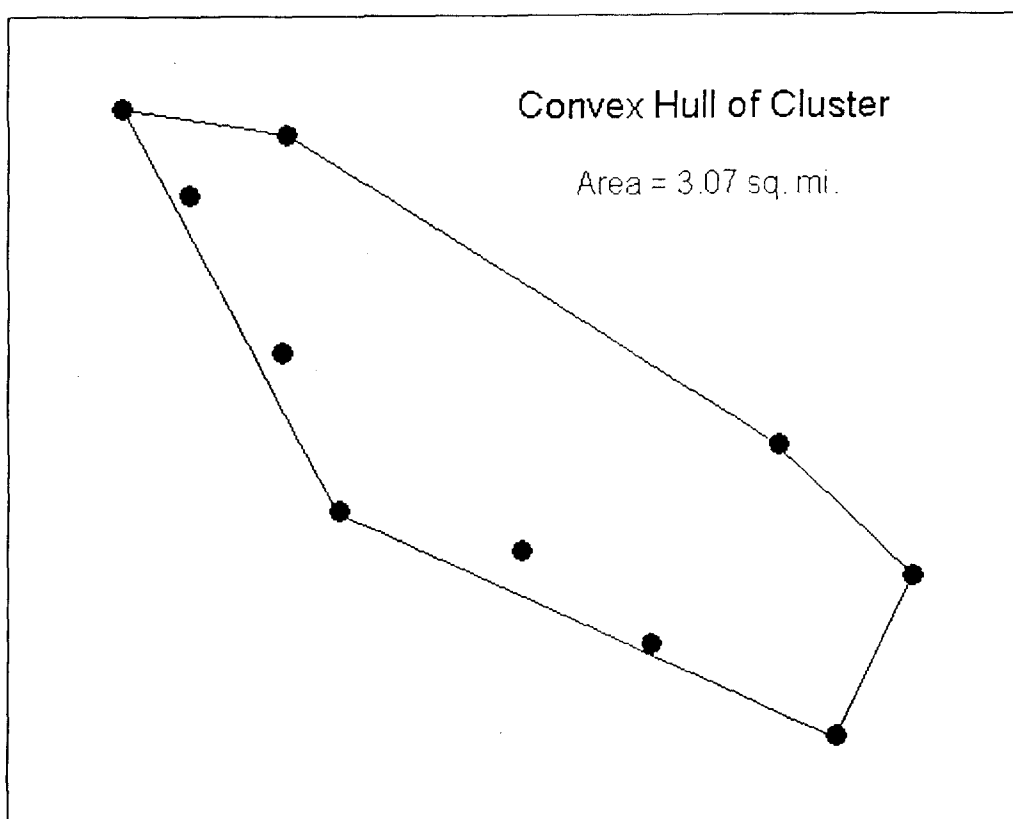
The Hatfield 5.0a Model groups a set of "actual" customer points into a *cluster*, according to a set of aggregation rules.



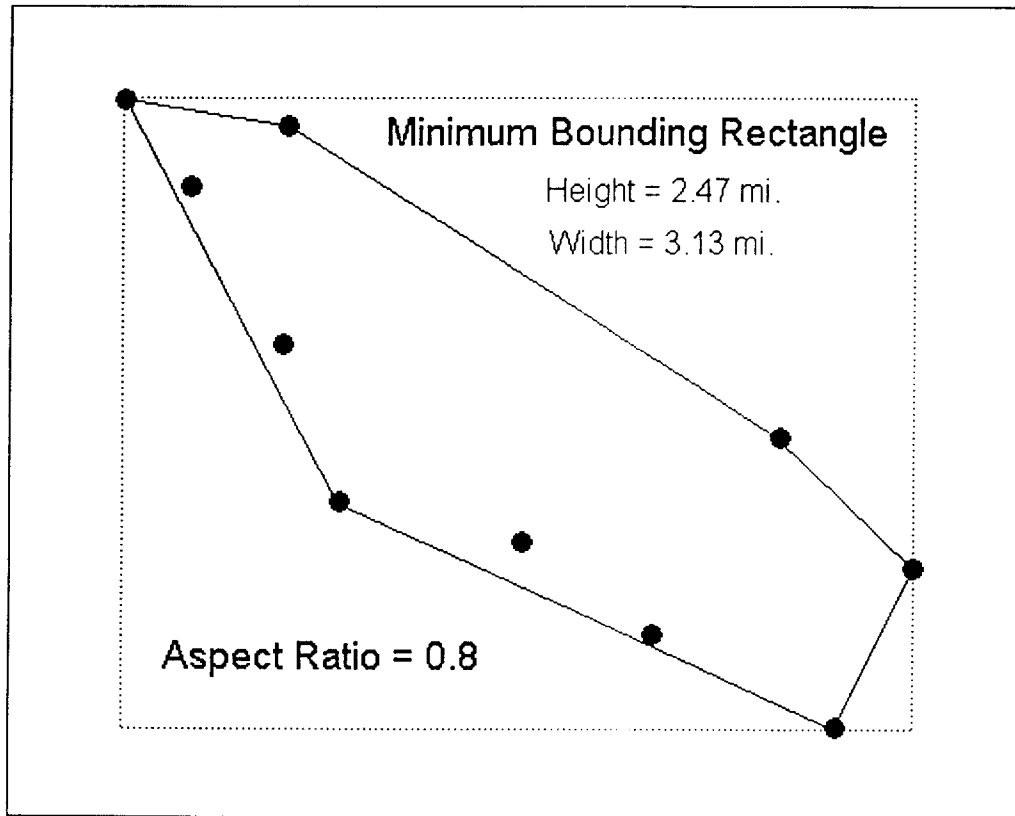
We have determined that the *minimum spanning tree* for these points – the mathematically shortest connection possible for these points – is 5.88 miles.



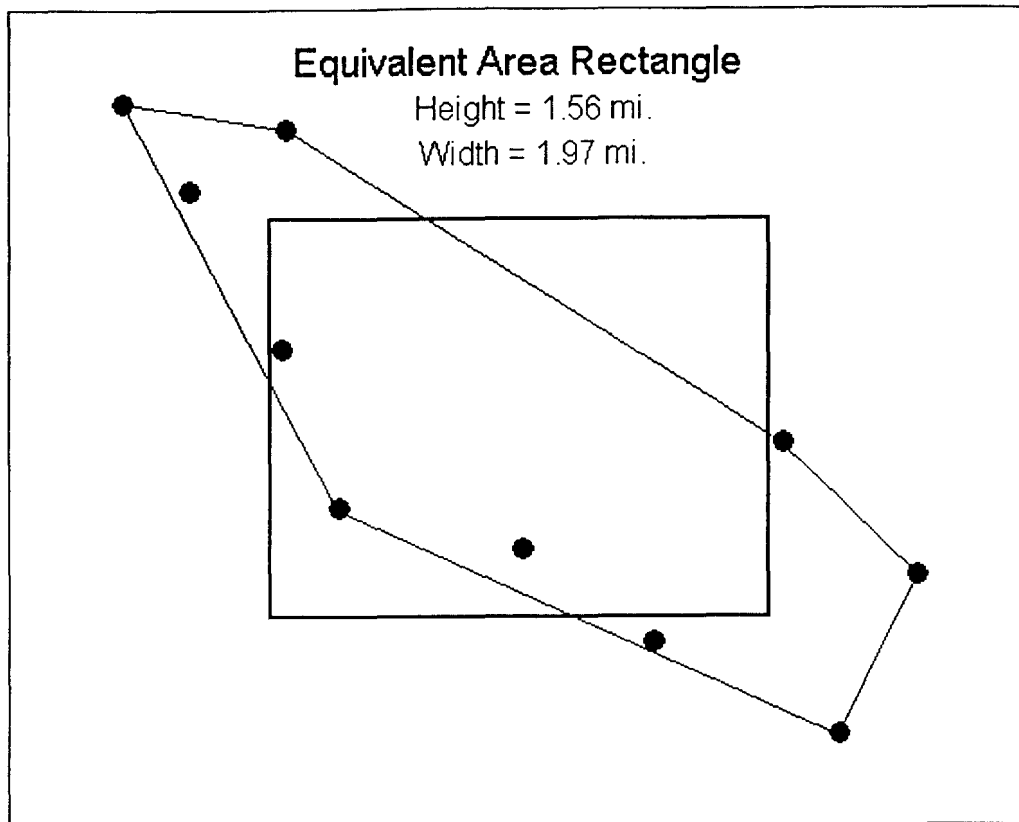
When Hatfield has determined the set of points that constitute a cluster, it logically draws a *convex hull* around those points, and determines its area.



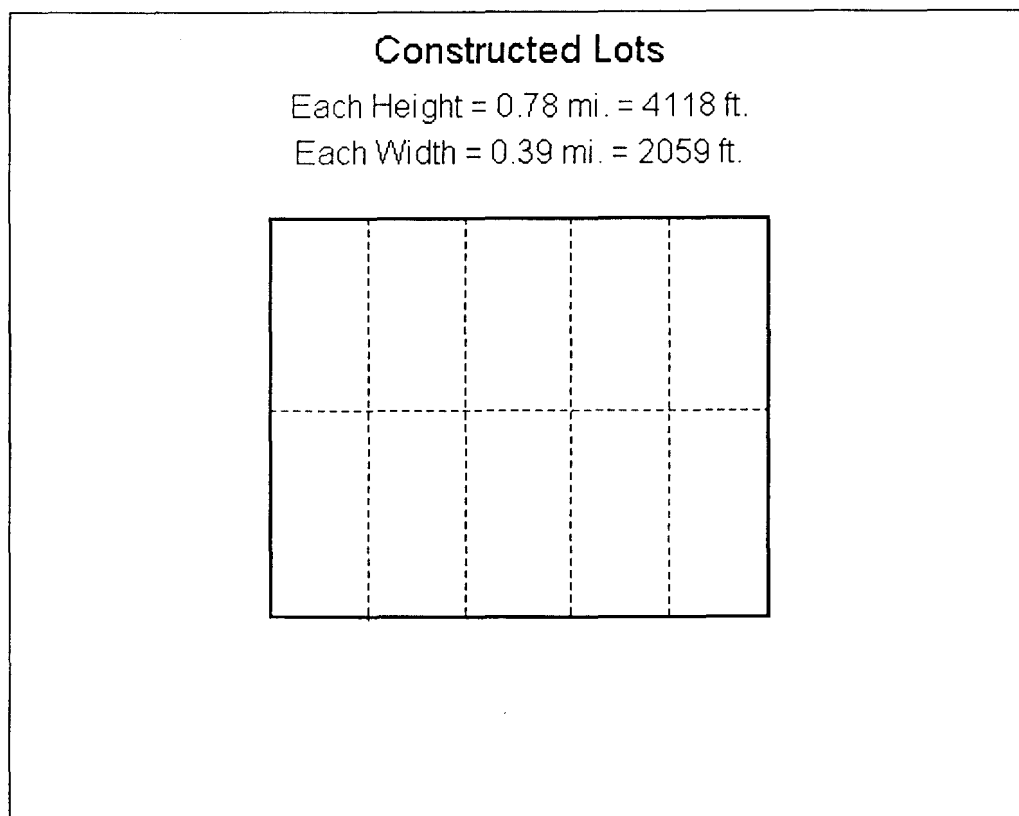
Hatfield then logically constructs a *minimum bounding rectangle* – oriented north-south-east-west – that exactly bounds the cluster's points. Hatfield then determines the *aspect ratio* of that rectangle (that is, the ratio of the rectangle's height to its width) ... in this case, 0.8.



Hatfield then constructs a *rectangle* with the above aspect ratio; the *size* of that rectangle is determined, of course, by its *area* ... and that area is set to be the *area of the convex hull* ... in this case, 3.07 square miles.



Hatfield then constructs *lots* within this constructed rectangle. Each lot is twice as high as it is wide.

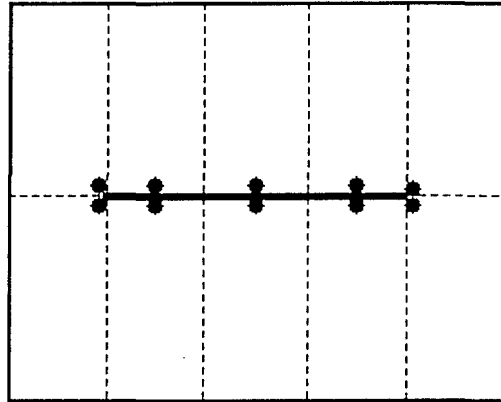


A *branch cable* is then constructed, and 150 ft. drops connect to the customers.

Cabling to Serve Customers

Branch Cable Length = 6177 ft.

10 Drops, each at 150 ft.



Total Cable Length = 7677 ft. = 1.45 mi.

Less than 1/4 of the Minimum Spanning Tree length!

But note how closely the customers are squeezed toward the branch cable. The arrangement is unrealistic, both from the standpoint of cable length *and* from the standpoint of area served.

Customer Area Served

Height = 300 ft.

Width = $106 + 6177 + 106$ ft. = 6389 ft.



Area Served = 1,916,700 sq. ft. = 0.0688 sq. mi.

But Actual Cluster Area = 3.07 sq. mi.

Area Modeled is 1/44 of Cluster Area

So, HOW BAD CAN THIS BE?

To what extent does the combined effect of:

- 1) converting the polygon into a rectangle (with identical area) and
- 2) building cable only to the point where the perimeter lots start
- 3) assuming all customers have drops 150 feet or less

cause the model to UNDERSTATE the amount of cable needed to transverse the ACTUAL distances between customers?

The following table shows a sample of several individual clusters (not wire centers) in Nevada (Nevada Bell territory).

The table gives an example of the amount of cable needed to reach all actual customer locations in the cluster. The locations do NOT include any outlier locations. The distance reported is only the distance between points that reside in the main clusters.

This length represents an approximation of the amount of distribution that the Hatfield Model (or any proxy model) should build in the course of laying out the network and determining the associated cost.

The table also shows the amount of actual distribution the Hatfield Model builds to each respective cluster (again, excluding outlier points).

Cluster Number	Absolute Minimum Distance Between Cluster Points (in feet)	Total Amount of Distribution Cable Built by Hatfield Model (in feet)
CHBTNV11.C003	23,500	7,900
IMLYNV12.C022	29,000	2,210
UPMDNVXF.C005	29,000	836
IMLYNV12.C015	38,000	2,089
DYTNNV11.C004	21,000	1,494
EMPRNV11.C004	21,500	5,093
EMPRNV11.C003	24,500	0

WHAT DOES THIS EVIDENCE EXPLAIN?

CONCLUSION #1: The Hatfield Sponsors' claim the placing surrogate points on the perimeters of CBs is a conservative approach (causing the model to overstate customer dispersion and therefore overstate required feet of plant) is completely false.

FACT: When points are placed in an (approximately) straight line, the area of the resulting polygon is miniscule and the converted rectangle with identical area distorts (understates) actual customer dispersion immensely.

CONCLUSION #2: This phenomenon has nothing to do with geocoding.

FACT: The understatement of plant does not depend on points being actual or surrogate. If a cluster is made up of 100% actual geocoded points and those points happen to be stretched out in a semi-linear fashion (i.e. along a road where geocoding places points), the same distortion will take place.

CONCLUSION #3: This also explains the significant differences in route mileage produced by the BCPM and the Hatfield Model for the same wire centers.

FACT: In many cases the BCPM estimates 10 times more distribution cable for a given wire center than the Hatfield Model does. Looking at only four clusters in the Imlay, NV wire center, we produce the same table:

Wire Center	Absolute Minimum Distance Between Cluster Points (in feet)	Total Amount of Distribution Cable Built by Hatfield Model (in feet)
4 Clusters in Imlay, NV (aggregated)	140,000	17,000